9-1998

Ab Initio Potential Energy Surface by Modified Shepard Interpolation: Application to the CH3+H2→CH4+H Reaction

Takeyuki Takata
University of Tokyo

Tetsuya Taketsugu
University of Tokyo

Kimihiko Hirao
University of Tokyo

Mark S. Gordon
Iowa State University, mgordon@iastate.edu

Ab Initio Potential Energy Surface by Modified Shepard Interpolation: Application to the CH3+H2→CH4+H Reaction

Abstract
An *ab initio*potential energy surface for the six-atom reactionCH3+H2→CH4+H was constructed, within C3v symmetry, by a modified Shepard interpolation method proposed recently by Collins et al. Selection of data points for the description of the potential energy surface was performed using both the Collins method and the dynamic reaction path (DRP) method. Although the DRP method is computationally more expensive, additional data points can be determined by just one simulation. Analyses of distributions of the data points, reaction probability, and errors in energy and energy gradients determined by the two different methods suggest a slight advantage for the DRP sampling in comparison with the iterative sampling.

Keywords
Potential energy surfaces, Surface reactions, Hydrogen reactions, Interpolation, Ab initio calculations

Disciplines
Chemistry

Comments
The following article appeared in *Journal of Chemical Physics* 109 (1998): 4281, and may be found at doi:10.1063/1.477032.

Ab initio potential energy surface by modified Shepard interpolation: Application to the CH 3 +H 2 →CH 4 +H reaction

Takeyuki Takata, Tetsuya Taketsugu, Kimihiko Hirao, and Mark S. Gordon

Citation: The Journal of Chemical Physics **109**, 4281 (1998); doi: 10.1063/1.477032
View online: http://dx.doi.org/10.1063/1.477032
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/109/11?ver=pdfcov
Published by the AIP Publishing

Ab initio potential energy surface by modified Shepard interpolation: Application to the $CH_3+H_2\rightarrow CH_4+H$ reaction

Takeyuki Takata, Tetsuya Taketsugu, and Kimihiko Hirao
Department of Applied Chemistry, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Mark S. Gordon
Department of Chemistry, Iowa State University, Ames, Iowa 50011

(Received 23 March 1998; accepted 12 June 1998)

An *ab initio* potential energy surface for the six-atom reaction $CH_3+H_2\rightarrow CH_4+H$ was constructed, within C_{3v} symmetry, by a modified Shepard interpolation method proposed recently by Collins *et al.* Selection of data points for the description of the potential energy surface was performed using both the Collins method and the dynamic reaction path (DRP) method. Although the DRP method is computationally more expensive, additional data points can be determined by just one simulation. Analyses of distributions of the data points, reaction probability, and errors in energy and energy gradients determined by the two different methods suggest a slight advantage for the DRP sampling in comparison with the iterative sampling. © *1998 American Institute of Physics.* [S0021-9606(98)30835-1]

I. INTRODUCTION

The construction of an accurate potential energy surface (PES) is an essential step in the analysis of the dynamics of a chemical reaction. The reliability of dynamic simulations is very sensitive to the quality of the PES employed. The PES, as a function of nuclear coordinates, can be determined, in principle, by solving the electronic Schrödinger equation for fixed nuclear arrangements[1] at some number of grid points in the appropriate configuration space. However, computational limitations have commonly necessitated the use of model potential energy functions (containing several fitting parameters) in molecular dynamics simulations. When the process includes a bond cleavage or bond formation, it is difficult to devise reliable potentials. As the number of atoms in the system of interest increases, a systematic determination of the PES becomes increasingly difficult, as does the subsequent process of fitting the resulting points to an analytic representation.

Recent developments in electronic structure theory, computational algorithms, and computer technology have made it possible to integrate molecular dynamics simulations into *ab initio* electronic structure codes. For example, the dynamic reaction path (DRP) method,[2–8] which was developed originally as an efficient method to locate the intrinsic reaction coordinate (IRC)[9–13] for elementary reactions, can be used to obtain a classical ''trajectory on-the-fly,'' without prior knowledge of the PES. The DRP method uses the energy gradient, obtained directly from electronic structure calculations, to determine atomic accelerations, velocities, and positions. Thus, this method explores the *ab initio* potential energy surface, although such *ab initio* trajectory calculations[14,15] require considerably more computational expense than traditional trajectory calculations.[16] The DRP method has recently been used to extract the characteristics of dynamical processes for several reactions.[4–8]

A promising method for incorporating *ab initio* potential energy surfaces into the dynamics of chemical reactions has recently been developed by Collins and co-workers,[17–25] and been applied to four-atom systems, including $NH+H_2\rightarrow NH_2+H$;[17] $OH+H_2\rightarrow H_2O+H$;[18–20,22] C_2H_2;[23] and $C(^3P)+H_3^+(^1A_1')$.[24] In their method, which has recently been extended to six-atom species,[25] the PES is constructed by a modified Shepard interpolation[26,27] in which the potential energy at a given configuration, **X**, is determined by using local Taylor expansions about N_d known data points. At those data points, the energy and derivatives in terms of nuclear coordinates are calculated by *ab initio* methods in advance. The total potential energy at **X** is evaluated as the sum over all N_d data points of their weighted Taylor series, where the weighting function becomes larger for data points closer to **X**. Methods for iteratively improving the PES using classical trajectory simulations have also been developed.[17,21] With this interpolation methodology, several groups have attempted to develop more efficient methods to construct the PES.[28–30]

For a given elementary reaction, there are usually some regions of configuration space that are more important than others. The most important region may be that within a vibrational amplitude of the minimum energy path (intrinsic reaction coordinate path, sometimes abbreviated as IRC); that is, the region of a minimum energy path which connects a reactant minimum, a transition state (TS), and a product minimum on the PES. In the Collins interpolation scheme, the initial PES is usually constructed in terms of data points distributed on the IRC[17,21] (referred to as the reaction path potential energy surface; RP-PES). Once the IRC is defined on the PES, the reaction path Hamiltonian[31] (RPH), which describes the motion along the IRC in terms of harmonic displacements in directions orthogonal to the reaction path, can be determined with a reasonable cost.[32] In addition to

being a common starting point for the study of reaction dynamics, the RPH provides important insights regarding the dynamics of a process of interest. If some amount of nuclear symmetry is conserved along the reaction path, Collins *et al.* make use of their symmetry-invariant reaction path potentials.[33]

Although the one-dimensional reaction path is extremely useful, the IRC has two geometrical weaknesses that can cause problems in dynamics calculations. These occur for sharply curved regions[5,34,35] and for unstable regions[36–40] on the reaction path. When a reaction path contains such regions, the significance of regions far from the IRC are expected to increase. It then becomes necessary to estimate the potential energy farther from the IRC than is ordinarily needed. Collins *et al.* employ classical trajectory calculations to select those data points which are far from the IRC but are important in dynamics; they are added to the set of original data points to improve the interpolated PES.[17,21]

In this paper, we apply the interpolation method of Collins *et al.*[17,21] to the construction of the PES for the six-atom reaction, $CH_3+H_2\rightarrow CH_4+H$, with 12 internal degrees of freedom. This reaction, including the nature of the IRC, has been extensively studied.[10,11,34,40–45] The IRC conserves C_{3v} symmetry throughout, and there are four totally symmetric (A_1) coordinates in this C_{3v} subspace. Since the reaction coordinate couples only with the orthogonal totally symmetric vibrations along the reaction path, the totally symmetric configuration space may be particularly important. The energy gradient on the IRC in orthogonal nontotally symmetric directions is zero, and the energy variation in those directions relative to the IRC can be described by an even function. Thus, as a first step in developing a global potential energy surface, we construct the PES in the totally symmetric configuration space (referred to as the totally symmetric PES), i.e., that part of the PES that may be expressed in terms of four totally symmetric coordinates in C_{3v} symmetry. The DRP method is employed for the selection of the data points to represent the PES (DRP sampling). Through application of the interpolation scheme to larger molecular systems, we investigate the applicability and limitations of the method.

II. INTERPOLATION SCHEME

We first summarize the modified Shepard interpolation method of Collins *et al.* for the construction and improvement of the PES,[17,21] with several additional modifications in the present study. In this interpolation scheme, the potential energy at some nuclear arrangement, **X**, can be expressed as the sum of contributions from several data points as

$$V^{\mathrm{MS}}(\mathbf{X})=\sum_{i=1}^{N_d} W_i(\mathbf{R})V_i(\mathbf{X}), \tag{1}$$

where **X** and **R** denote $3N$ Cartesian and $3N-6$ internal coordinates, respectively, for an N atom system, $W_i(\mathbf{R})$ denotes a weight function for the ith data point, $V_i(\mathbf{X})$ is a Taylor expansion function about $\mathbf{X}^{(i)}$, the Cartesian coordinates for the ith data point, and N_d is the number of data points. Note that Cartesian coordinates are used to calculate potential energy values,[46] while the weight function is based on internal coordinates. This modification of the original scheme is made, because the potential energy can be more easily calculated in terms of Cartesian coordinates in the present case, and the weight function should depend on only the molecular internal structure. For the Taylor expansion about data points, we employ the second order series as

$$V_i(\mathbf{X})=V^{(i)}+\sum_{j=1}^{3N}(X_j-X_j^{(i)})g_j^{(i)}+\frac{1}{2}\sum_{j,k=1}^{3N} f_{jk}^{(i)}(X_j-X_j^{(i)})(X_k-X_k^{(i)}), \tag{2}$$

where $V^{(i)}$, $g_j^{(i)}$, and $f_{jk}^{(i)}$ are the potential energy, the jth component of the gradient vector, and the jkth component of the Hessian matrix, respectively, calculated by *ab initio* electronic structure calculations at $\mathbf{X}^{(i)}$. The utility of higher order derivatives (third and fourth) was investigated by Collins *et al.*,[19] and it was concluded that, without an efficient means for calculating analytic third (or fourth) derivatives, second order derivatives provide the most cost-effective means of constructing a global PES by interpolation. The weight function $W_i(\mathbf{R})$, which approaches 1 as $\mathbf{R}\rightarrow\mathbf{R}^{(i)}$ and approaches 0 as $|\mathbf{R}-\mathbf{R}^{(i)}|\rightarrow\infty$, is given by

$$W_i(\mathbf{R})=\frac{w_i(\mathbf{R})}{\Sigma_{k=1}^{N_d} w_k(\mathbf{R})}, \tag{3}$$

$$w_i(\mathbf{R})=\left\{\sum_{j}^{3N-6}(R_j-R_j^{(i)})^2\right\}^{-p}. \tag{4}$$

Equation (3) assures the normalization of the weight function $(\Sigma_{i=1}^{N_d} W_i(\mathbf{R})=1)$. The parameter $2p$ in Eq. (4) must be greater than the larger of $3N-6$ and the order of the Taylor series expansion to guarantee convergence of the potential $V^{\mathrm{MS}}(\mathbf{X})$.[18]

The simplest PES for a given elementary reaction is the reaction path (RP)-PES. Within the context of the interpolation scheme, the RP-PES can be constructed in terms of data points that are evenly spaced along the IRC. The utility of the RP-PES is based on the significance of the (static) reaction path. Since a chemical reaction is a dynamical process, it is important to include points in dynamically significant regions in the interpolation set. For this purpose, Collins *et al.*[17–21] devised the following procedures: First, the RP-PES is constructed in terms of data points on the IRC; then classical trajectory calculations are carried out on this PES. Weights are assigned for selected points on these classical trajectories. These weights are large in regions that are (1) sampled by many trajectories and (2) far from data points that are already included. The former condition indicates the dynamical significance of that nuclear arrangement. The following weight function satisfying these two conditions was introduced by Collins *et al.*,[17,21]

$$h[\mathbf{R}(j)]=\frac{1}{N_T-1}\frac{\Sigma_{n(\neq j)}^{N_T} w_n[\mathbf{R}(j)]}{\Sigma_k^{N_d} w_k[\mathbf{R}(j)]}, \tag{5}$$

where $\mathbf{R}(j)$ denotes a set of $3N-6$ internal coordinates at the jth selected point from a trajectory, N_T is the number of selected points, and N_d is the number of data points for interpolation.[47] The functional form of $w_k(\mathbf{R})$ is given in Eq. (4). The point giving the maximum weight is added to the group of data points. Then, the contribution from this new point moves from the numerator to the denominator in Eq. (5), and the weights in Eq. (5) are recalculated for the remaining trajectory points that have been selected. This selection of data points is repeated until n_T new data points are chosen (n_T is a parameter). Finally, *ab initio* electronic structure calculations are performed to determine the energy, energy gradient, and Hessian matrix at these new data points, resulting in an improved interpolated PES. Classical trajectory calculations are again carried out on this revised PES, then n_T additional points are selected again according to Eq. (5), and implemented to a group of data points.

The procedure described above for improving the PES is continued until one or more physical quantities that are determined by the PES (e.g., the reaction probability or errors in the potential energy or energy derivatives) have converged. The reaction probability can be calculated as the fraction of trajectories that reach the product side. The error in the PES itself can be estimated by the averaged deviations of the energy and energy gradient at several selected reference points,

$$\mathrm{ERR}_V=\sum_i^M |V^{\mathrm{MS}}(\mathbf{X}_i)-V^{\text{ab initio}}(\mathbf{X}_i)|/M, \tag{6}$$

$$\mathrm{ERR}_G=\sum_i^M |G^{\mathrm{MS}}(\mathbf{X}_i)-G^{\text{ab initio}}(\mathbf{X}_i)|/M, \tag{7}$$

where V^{MS} and $V^{\text{ab initio}}$ denote the potential energy at $\mathbf{X}_i$ (ith reference point) estimated by the (modified Shepard) interpolation method and the *ab initio* electronic structure calculation, respectively; G^{MS} and $G^{\text{ab initio}}$ are the corresponding magnitudes of the energy gradient at $\mathbf{X}_i$, and M is the total number of reference points used for the error analysis.

An alternative method is proposed in this work for the selection of dynamically important points to be used for the interpolation process. The original Collins method[17,21] improves the PES iteratively based on classical trajectories, which run on the updated interpolated PES. In the initial stages, the trajectories will contain some level of inaccuracy; as the PES converges to the correct result, the trajectories also converge to the correct ones. This approach may be referred to as iterative sampling. As an alternative for obtaining a correct *ab initio* trajectory, one can use the dynamic reaction path (DRP) method,[2–8] in which a trajectory is integrated according to Newton's equation of motion by using the potential energy and energy gradient calculated directly by *ab initio* electronic structure methods. In this method, the only inherent approximation is the computational level of *ab initio* theory employed. Of course, the tradeoff is the expense of a directly calculated *ab initio* trajectory which limits the number of trajectories that one can practically calculate. The DRP method can be used to obtain correct reference values for physical quantities (e.g., the reaction probability; the trace of a trajectory) to check the convergence of the interpolated PES. This DRP method can be used to directly determine dynamically important points with just one trial (i.e., DRP sampling). In the present study, we carry out DRP sampling for the selection of the data points, and compare the PES obtained in this manner with that obtained by iterative sampling.

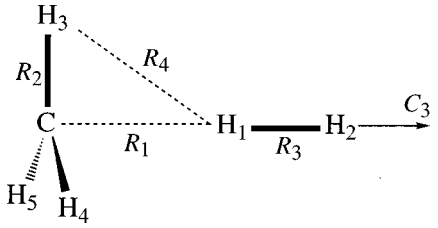

FIG. 1. Four total symmetry coordinates, R_1-R_4, which describe the reaction, $CH_3+H_2\rightarrow CH_4+H$, in C_{3v} subspace. R_2 and R_4 represent three equivalent atomic distances of CH and HH, respectively.

III. REACTION SYSTEM: $CH_3+H_2\rightarrow CH_4+H$

The modified Shepard interpolation method[17,21] is applied to the six-atom reaction, $CH_3+H_2\rightarrow CH_4+H$. As described in the Introduction, this IRC retains C_{3v} symmetry throughout, thus the totally symmetric PES for this reaction can be described in terms of four totally symmetric coordinates.[34] In the present study, this four-dimensional totally symmetric PES is constructed as the first step in building the entire 12-dimensional PES (the totally symmetric PES may be the most important region for the reaction dynamics). Note that a classical trajectory also retains the spatial symmetry determined by the initial conditions (i.e., atomic positions and velocities). Figure 1 shows the four totally symmetric coordinates, R_1-R_4, employed in the present study. The positions of H_3, H_4, and H_5 change so that the system retains C_{3v} symmetry, thus R_2 and R_4 each represent three equivalent atomic distances in C_{3v} symmetry. Then, molecular structures in the C_{3v} subspace can be expressed by R_1-R_4. These internal coordinates are used in the weight function in Eq. (4) for the interpolation. Since this method requires a large number of *ab initio* electronic structure calculations, and the main intent of this work is to demonstrate the utility of the method, the unrestricted Hartree–Fock (UHF) method is employed with the 6-31G(d,p) basis set[48] using the GAMESS program.[49]

Figure 2 shows variations in (a) energy and (b) geometrical parameters along the IRC. The activation energy for this reaction is estimated as 21.4 kcal/mol, and the energy of products, CH_4+H, is calculated as -2.4 kcal/mol relative to that of reactants, CH_3+H_2. Figure 2(b) indicates that this reaction can be characterized by changes in R_1 ($C\cdots H_1$) and R_3 ($H_1\cdots H_2$). R_2 remains nearly constant ($\sim$1.08 Å) throughout, while changes in R_4 reflect those in R_1, so R_4 can be expressed by a linear function of R_1 along the IRC. In the following, we discuss the quality of the PES based on two-dimensional sections described in terms of R_1 and R_3, with R_2 fixed at 1.08 Å. R_4 is determined by the aforementioned linear function of R_1. Figure 3 shows the *ab initio* PES in the vicinity of the transition state, constructed in

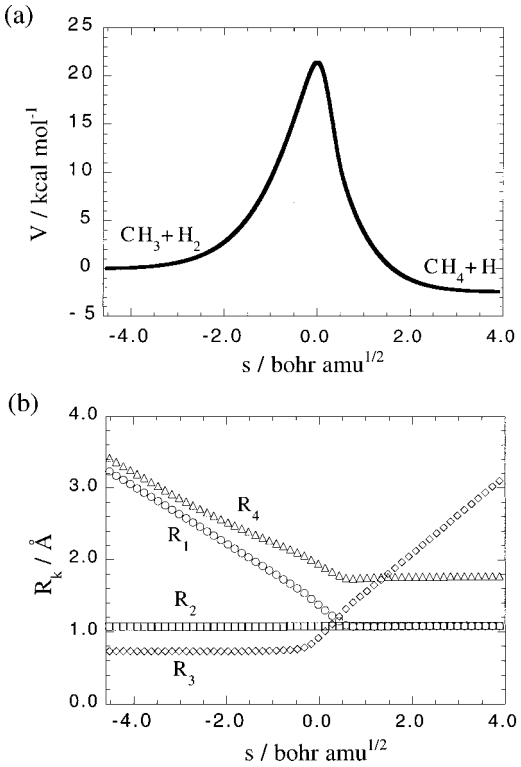

FIG. 2. Variations of (a) energy and (b) geometrical parameters, R_1–R_4, along the IRC ($CH_3+H_2 \rightarrow CH_4+H$).

terms of *ab initio* (UHF) energies at 400 grid points. The energy difference between the contour lines is 5 and the maximum contour is set to 40 kcal/mol. This *ab initio* PES will be used as a reference to check the quality of the interpolated PES.

To integrate a trajectory, we employ the algorithm proposed by Stewart *et al.* for the DRP method.[2] The time step is fixed at 0.1 fs. In the initial stages of trajectory calculations, quantization of vibrational motions is taken into account as in quasiclassical trajectory simulations. Within C_{3v} subspace, there are three internal totally symmetric normal vibrational modes in the reactants (CH_3+H_2): CH_3 stretching (3260.0); CH_3 umbrella (339.6); HH stretching (4635.1 cm^{-1}). The remaining totally symmetric mode corresponds to the collisional motion of CH_3 and H_2 in which C_{3v} symmetry is retained. Zero-point energy is assigned to the three internal vibrational modes, and 35 kcal/mol is provided to the collisional mode. The latter is based on preliminary calculations and the observation that too high a collision energy leads to a reaction probability of 100%, while too low an energy leads to no reaction. Within the harmonic approximation, the ith normal coordinate Q_i and its conjugate momentum P_i in the vibrational ground states are

$$Q_i = \sqrt{h\nu_i/\lambda_i}\cos\omega_i, \tag{8}$$

$$P_i = \sqrt{h\nu_i}\sin\omega_i, \tag{9}$$

where ν_i and λ_i denote the vibrational frequency and force constant of the ith normal mode, respectively, ω_i denotes a phase angle, and h is Planck's constant. Four phase angles, 0, $\pi/2$, π, $3\pi/2$, are considered. This leads to 64 ($=4^3$) initial

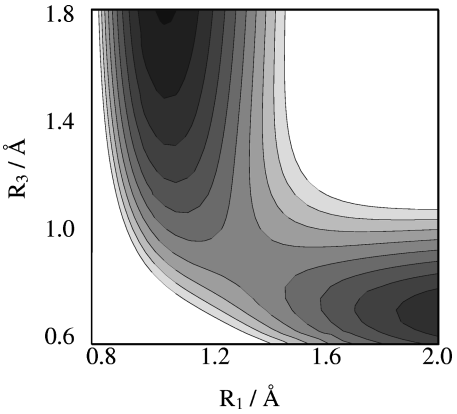

FIG. 3. *Ab initio* PES around a TS region, constructed in terms of UHF energies at 400 grid points.

conditions for the trajectory calculations. Trajectories are initiated in the region of $R_1 \sim 2.5$ Å. To provide a range of initial conditions, the initial value of R_1 is taken as

$$R_1 = 2.5 + 0.1 \times \sin\left(\frac{2n\pi}{4}\right), \tag{10}$$

where n refers to 1 of the 64 trajectories. The trajectory calculation is continued for 60 fs or until H_2 is scattered from CH_3; that is, if R_1 decreases from ~2.5 to below 1.6 Å and then back to distances that are larger than 1.6 Å. The criterion for reaction to CH_4+H is taken to be $R_3 = 1.5$ Å.

The unnormalized weight function $w_i(\mathbf{R})$ in Eq. (4) includes a parameter, p, which must be greater than half the value of the larger of (the number of internal degrees of freedom/the order of the Taylor series expansion).[18] As p is increased, the weight for the closest point in a normalized weight function [Eq. (3)] increases since $R_j - R_j^{(i)} < 1$. With a sufficiently large value of p, one obtains a smoothly varying PES. An extensive preliminary survey suggested that $p \geqslant 4$. So, p is set to 4 in this study. In order to choose new data points and conserve CPU time, the weight given in Eq. (5) is assigned for one-seventh of all the points on trajectories. In the iterative sampling, the optimal number of data points added in one cycle (n_T) was determined to be 30 through several trial and error calculations. By repeating this process ten times, 300 data points are added into the group of data points for the interpolation.

IV. RESULTS AND DISCUSSION

In the interpolation scheme for the PES, the initial PES, i.e., the reaction path PES (RP-PES) constructed from data points evenly scattered along the IRC, may be the most primitive. With this RP-PES, the energy for a given point can be calculated by the weighted average of the second order Taylor series about the respective data points as indicated by Eqs. (1) and (2). In the present study, 39 points on the IRC are chosen as data points: 1 corresponding to the TS, 19 points from the reactant side, and 19 points from the product side. Figure 4 shows (a) positions of these data points along the IRC and (b) the calculated RP-PES in the (R_1, R_3) plane (using interpolated energies at 3600 grid

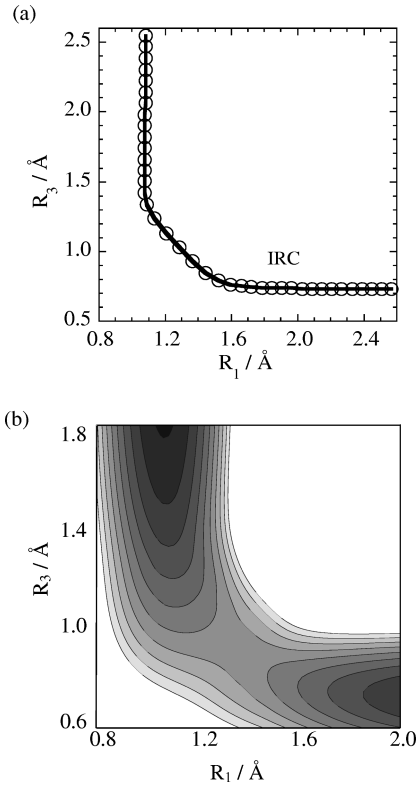

FIG. 4. (a) Positions of the initial data points evenly spaced along the IRC and (b) the calculated RP-PES, in a (R_1, R_3) plane.

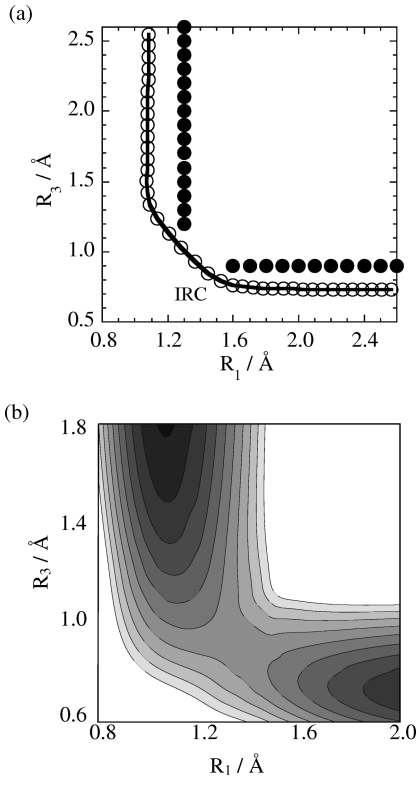

FIG. 5. (a) Positions of the initial data points (white circles for those on the IRC and black circles for those in the concave side of the IRC) and (b) the interpolated PES calculated with these data points.

points). In this figure, R_2 is fixed at 1.08 Å while R_4 is determined as a linear function of R_1 (see Sec. III). One can see from Fig. 4(a) that no significant molecular deformation (abstraction of H) occurs until (approaching from the right in the figure) R_1 decreases to ~1.6 Å. The structural transformation continues until R_3 exceeds about 1.3 Å [see also Fig. 2(a)]. In Fig. 4(b), the focus is on the PES around this important region close to the transition state. The energy differences between the contour lines is the same as that in Fig. 3. Since the harmonic approximation has been used for the directions orthogonal to the reaction path, the energy variation on both sides of the IRC looks symmetric relative to the IRC, in contrast with the PES in Fig. 3.

The energy variation in bond stretching directions is more appropriately expressed by a Morse function than by a quadratic (harmonic) function. In the Morse function, the energy increase along the positive coordinate (toward bond dissociation) is softened, while it becomes harder along the negative coordinate. So, for a given energy, the system is more easily displaced (i.e., is more anharmonic) in the positive than in the negative direction. The reaction system in the present study has four totally symmetric coordinates, each of which may be classified as a bondingtype coordinate. Since the reaction path is characterized explicitly by two of these coordinates, R_1 and R_3, as illustrated in Fig. 2, 28 data points were added to the positive sections of $R_1(=1.3\,\text{Å})$. and $R_3(=0.9\,\text{Å})$ relative to the IRC in preparing the initial PES [indicated by the black circles in Fig. 5(a)]. These additional data points are approximately on the reaction plane determined by the reaction path tangent and curvature vectors, which are frequently important in dynamics.[5,34,35] Figure 5 shows (a) positions of the initial data points (white circles for those on the IRC and black circles for those in the concave side of the IRC; 67 in total) and (b) the interpolated PES calculated with these data points. As illustrated in Fig. 5(b), the energy on the concave side of the IRC is lowered considerably in comparison with that in Fig. 4(b) (RP-PES). In the following, the PES in Fig. 5(b) is employed as the initial PES to be improved.

Figure 6 shows (a) traces of 64 trajectories calculated on the initial PES with 67 data points (white circles), and (b) traces of 64 DRP trajectories calculated with the same initial conditions. Note that the simulations presented here are actually carried out on a four-dimensional PES, the four dimensions corresponding to the four totally symmetric internal coordinates. The DRP simulations are exact on the UHF/6-31G(d,p) PES. This computational level is also employed for the *ab initio* molecular orbital (MO) energies and derivatives used here to construct the interpolated PES. Thus, the DRP simulations can be used as a reference to test the quality of the approximate PES employed. At first glance, these two sets of trajectories exhibit similar behavior, but there are some interesting differences: DRP trajectories can deviate more extensively relative to the IRC in the product (left) side. In addition, two trajectories return across the concave side of the TS toward the reactant side in Fig. 6(a), while only one trajectory returns in this manner in Fig. 6(b). Overall, more trajectories seem to return to the reactant side in Fig. 6(b), judging from the density difference of trajectories. This last point can be verified by estimating the reaction

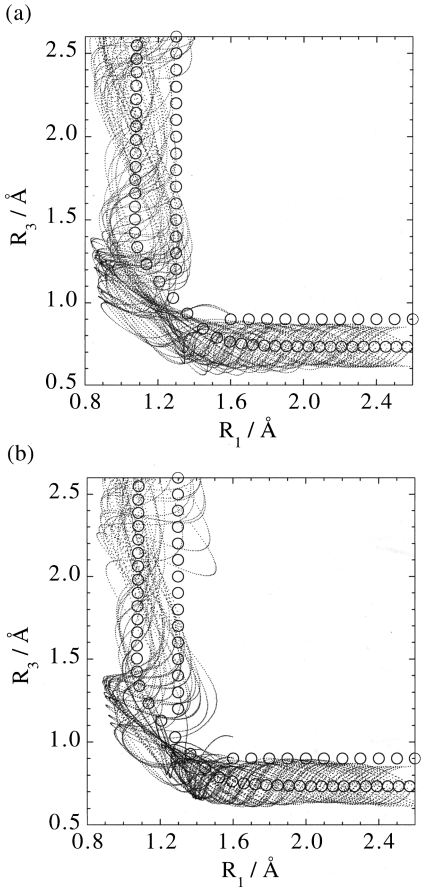

FIG. 6. (a) Traces of 64 trajectories calculated on the initial PES with 67 data points of white circles and (b) DRP trajectories calculated with the same initial conditions.

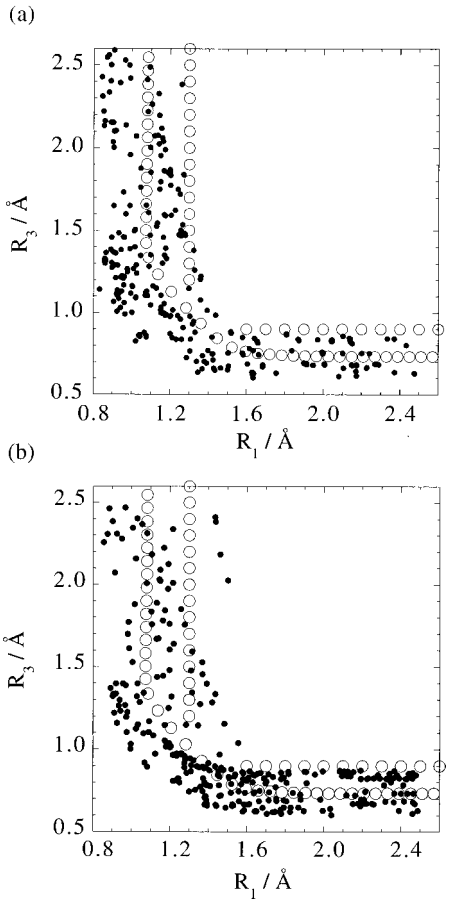

FIG. 7. Distributions of initial (white circle) and additional (black circle) data points for the interpolation: (a) the iterative sampling; (b) the DRP sampling.

probability in the respective cases: 88% in Fig. 6(a), 75% in Fig. 6(b).[50] These differences may be due to differences in geometric features of the PES near the TS region.

Next, we proceed to improved potential energy surfaces, in which the additional data points selected either by iterative sampling or by DRP sampling have been included. Using iterative sampling, 30 data points are added by one simulation (64 trajectory calculations on a given PES); this procedure is repeated ten times, resulting in a total of 300 additional data points. So, in this sampling, the PES is improved step by step through iterative trajectory calculations. On the other hand, in the DRP sampling, 300 data points for the interpolation can be added through just one simulation, since the DRP trajectory runs on the exact PES for this level of theory (of course, both sets of calculations are limited by the approximations in the *ab initio* MO method employed). The obvious disadvantage of the DRP sampling is the computational cost.

The selection of data points was performed using the weight function in Eq. (5) for both methods. Figure 7 shows distributions of initial (white circle) and additional (black circle) data points selected by (a) the iterative sampling and (b) the DRP sampling. For both methods, many data points are observed on the convex side of the IRC around the TS. This is due to the centrifugal force acting on the reaction system in that region due to the curvature of the IRC. Comparison of Figs. 7(a) and 7(b) shows that the data points are distributed more or less uniformly along the IRC in the iterative sampling procedure, while the data points accumulate more on the reactant (right) side in the DRP sampling. Those distributions reflect the distributions of trajectories shown in Fig. 6.

Figure 8 shows the improved PES with the additional data points selected by (a) the iterative sampling and (b) the DRP sampling, corresponding to Figs. 7(a) and 7(b), respectively. In comparison with the *ab initio* PES in Fig. 3, they have somewhat ragged regions; the additional data points on the convex side of the IRC lower the potential energy around $(R_1, R_3) = (1.0, 1.0)$. Figures 9(a) and 9(b) show traces of 64 trajectories calculated on the improved PES with 367 data points, corresponding to Figs. 8(a) and 8(b), respectively. In comparison with Fig. 6(a), they seem to approach more correct behavior, especially on the product side.

To check the quality of the PES quantitatively, the reaction probability and errors in the potential energy [Eq. (6)] and energy gradient [Eq. (7)] were calculated using the two potential energy surfaces. To determine how the interpolated PES converges as additional data points are added, the reaction probability and errors for the DRP-sampling-PES were calculated at every 20th additional data point. Figure 10 shows variations in the reaction probability for the iterative sampling and the DRP sampling. The dotted line corresponds

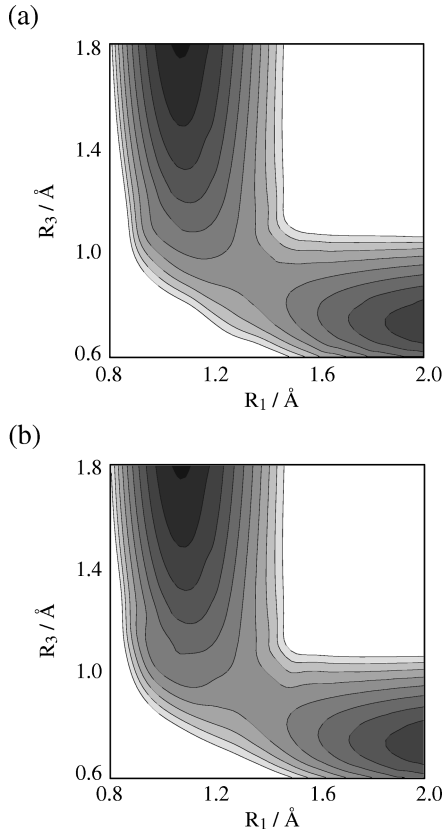

FIG. 8. The improved PES interpolated by the initial data points plus additional ones selected by (a) the iterative sampling and (b) the DRP sampling.

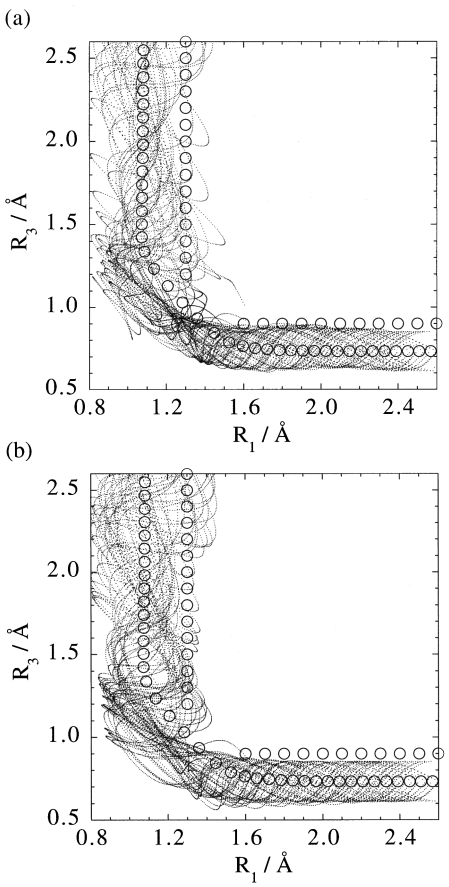

FIG. 9. Traces of 64 trajectories calculated on the improved PES with 67 + 300 data points selected by (a) the iterative sampling and (b) the DRP sampling.

to the exact value (75%) for the UHF/6-31G(d,p) PES estimated by DRP simulations. Due to the small number of additional data points, the reaction probability decreases rapidly from 88% to about 70% in both sampling methods. They show fluctuations around 70% for a while, then converge near 75% with more than 200 data points. As reference points for estimating errors in the energy and gradient from *ab initio* values [Eqs. (6) and (7)], 448 data points were selected from points on the DRP trajectories at random. Figure 11 shows changes in these errors in (a) energy and (b) energy gradient against the number of additional data points. Both figures indicate a slight advantage of the DRP sampling in comparison with the iterative sampling. The errors decrease rapidly for both methods after adding the first group of data points, then undergo relatively small changes as the number of data points is increased further.

V. CONCLUSION

In the present paper, we have applied the modified Shepard interpolation method proposed by Collins *et al.* to the construction of the PES in a polyatomic reaction including six atoms, $CH_3+H_2 \rightarrow CH_4+H$. As a starting point for an eventual study of the entire PES, the focus is given to the totally symmetric (four-dimensional) PES in C_{3v} symmetry, which is conserved along the IRC. The weight function is described in terms of four totally symmetric (internal) coordinates, while the potential energy is calculated based on a second order Taylor series in Cartesian coordinates. In this methodology, starting from the reaction path PES, the PES is improved by adding data points chosen using both the DRP sampling method, as well as the iterative method developed by Collins and co-workers. Although the DRP method requires a greater computational cost, the additional data points

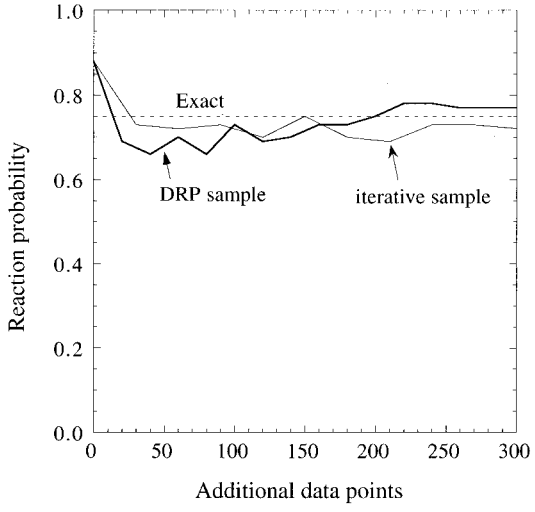

FIG. 10. Reaction probability as a function of number of additional data points by the iterative sampling and by the DRP sampling. The exact value is calculated by the DRP simulation.

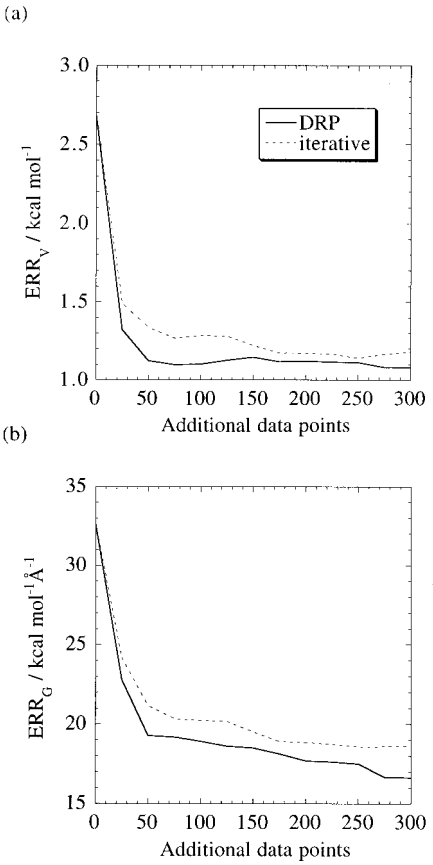

FIG. 11. Error of (a) energy (ERR_V) and (b) energy gradient (ERR_G) against the number of additional data points by the iterative sampling and by the DRP sampling.

can be determined by just one simulation. Comparison of the data points determined by the two different methods suggests that, although their convergence of the reaction probability is similar to each other, the distributions of data points and the convergence in errors of energy and energy gradient seem to be a little better in the DRP sampling than in the iterative sampling. Since the DRP method is generally applicable to molecular systems of any size, the present study suggests that the DRP method should be an effective partner to the interpolation method for the construction of potential energy surfaces.

ACKNOWLEDGMENTS

We are grateful to Dr. Nobuo Tajima, Mr. Kenichi Nakayama, and Mr. Tomohiro Hashimoto for helpful suggestions on calculations of standard deviations of reaction probability. T.T., T.T., and K.H. were supported by Grant-in-Aid for Scientific Research on Priority Areas from Ministry of Education, Science and Culture in Japan, and M.S.G. was supported by grants from the National Science Foundation and the Air Force Office of Scientific Research.

[1] M. Born and R. Oppenheimer, Ann. Phys. **84**, 457 (1927).
[2] J. J. P. Stewart, L. P. Davis, and L. W. Burggraf, J. Comput. Chem. **8**, 1117 (1987).
[3] S. A. Maluendes and M. Dupuis, J. Chem. Phys. **93**, 5902 (1990).
[4] T. Taketsugu and M. S. Gordon, J. Phys. Chem. **99**, 8462,14597 (1995).
[5] T. Taketsugu and M. S. Gordon, J. Chem. Phys. **103**, 10042 (1995).
[6] M. S. Gordon, G. Chaban, and T. Taketsugu, J. Phys. Chem. **100**, 11512 (1996).
[7] T. Taketsugu and T. Hirano, J. Chem. Phys. **104**, 6081 (1996).
[8] T. Taketsugu, T. Yanai, K. Hirao, and M. S. Gordon, J. Mol. Struct.: THEOCHEM (in press).
[9] D. G. Truhlar, R. Steckler, and M. S. Gordon, Chem. Rev. **87**, 217 (1987).
[10] B. C. Garrett, M. Redmon, R. Steckler, D. G. Truhlar, K. K. Baldridge, D. Bartol, M. W. Schmidt, and M. S. Gordon, J. Phys. Chem. **92**, 1476 (1988).
[11] K. K. Baldridge, M. S. Gordon, D. G. Truhlar, and R. Steckler, J. Phys. Chem. **93**, 5107 (1989).
[12] D. G. Truhlar and M. S. Gordon, Science **249**, 491 (1990).
[13] *The Reaction Path in Chemistry: Current Approaches and Perspectives*, edited by D. Heidrich (Kluwer, Netherlands, 1995).
[14] (a) T. Helgaker, E. Uggerund, and H. J. A. Jensen, Chem. Phys. Lett. **173**, 145 (1990); (b) E. Uggerund and T. Helgaker, J. Am. Chem. Soc. **114**, 4265 (1992).
[15] W. Chen, W. L. Hase, and H. B. Schlegel, Chem. Phys. Lett. **228**, 436 (1994).
[16] F. B. Brown and D. G. Truhlar, Chem. Phys. Lett. **117**, 307 (1985).
[17] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).
[18] M. J. T. Jordan, K. C. Thompson, and M. A. Collins, J. Chem. Phys. **102**, 5647 (1995).
[19] M. J. T. Jordan, K. C. Thompson, and M. A. Collins, J. Chem. Phys. **103**, 9669 (1995).
[20] M. J. T. Jordan and M. A. Collins, J. Chem. Phys. **104**, 4600 (1996).
[21] M. A. Collins, in *New Method in Computational Quantum Mechanics*, edited by I. Prigogine and S. A. Rice (Wiley, New York, 1996), p. 389.
[22] K. C. Thompson and M. A. Collins, J. Chem. Soc., Faraday Trans. **93**, 871 (1997).
[23] K. C. Thompson, M. J. T. Jordan, and M. A. Collins, J. Chem. Phys. **108**, 564 (1998).
[24] R. P. A. Bettens and M. A. Collins, J. Chem. Phys. **108**, 2424 (1998).
[25] Private communications.
[26] R. Farwig, in *Algorithms For Approximation*, edited by J. C. Mason and M. G. Cox (Clarendon, Oxford, 1987), p. 194.
[27] P. Lancaster and K. Salkauskas, *Curve and Surface Fitting, An Introduction* (Academic, London, 1986).
[28] K. A. Nguyen, I. Rossi, and D. G. Truhlar, J. Chem. Phys. **103**, 5522 (1995).
[29] Y. M. Rhee, T. G. Lee, S. C. Park, and M. S. Kim, J. Chem. Phys. **106**, 1003 (1997).
[30] T. Ishida and G. C. Schatz, J. Chem. Phys. **107**, 3558 (1997).
[31] W. H. Miller, N. C. Handy, and J. E. Adams, J. Chem. Phys. **72**, 99 (1980).
[32] D. G. Truhlar and B. C. Garrett, Acc. Chem. Res. **13**, 440 (1980); Annu. Rev. Phys. Chem. **35**, 189 (1984); D. G. Truhlar, R. Steckler, and M. S. Gordon, Chem. Rev. **87**, 217 (1987).
[33] J. Ischtwan and M. A. Collins, J. Chem. Phys. **94**, 7084 (1991).
[34] T. Taketsugu and M. S. Gordon, J. Chem. Phys. **104**, 2834 (1996).
[35] T. Taketsugu and K. Hirao, J. Chem. Phys. **107**, 10506 (1997).
[36] T. Taketsugu and T. Hirano, J. Chem. Phys. **99**, 9806 (1993).
[37] T. Taketsugu and T. Hirano, J. Mol. Struct.: THEOCHEM **310**, 169 (1994).
[38] T. Taketsugu, N. Tajima, and K. Hirao, J. Chem. Phys. **105**, 1933 (1996).
[39] T. Yanai, T. Taketsugu, and K. Hirao, J. Chem. Phys. **107**, 1137 (1997).
[40] K. Yamashita and T. Yamabe, Int. J. Quantum Chem., Symp. **17**, 177 (1983).
[41] R. Steckler, K. J. Dykema, F. B. Brown, G. C. Hancock, D. G. Truhlar, and T. Valencich, J. Chem. Phys. **87**, 7024 (1987).
[42] T. Joseph, R. Steckler, and D. G. Truhlar, J. Chem. Phys. **87**, 7036 (1987).
[43] J. A. Boatz and M. S. Gordon, J. Phys. Chem. **93**, 5774 (1989).
[44] T. N. Truong, J. Chem. Phys. **100**, 8014 (1994).
[45] M. J. T. Jordan and R. G. Gilbert, J. Chem. Phys. **102**, 5669 (1995).
[46] The major disadvantage of expanding a PES in Cartesian coordinates is that a general expression for the potential energy will not be invariant to the translation, rotation, and inversion of the molecule: See Ref. 23. However, the present study is restricted to the totally symmetric PES in C_{3v} symmetry, so the coordinate axes (x,y,z) need not be reoriented.
[47] Collins *et al.* introduced two weight functions which are an alternative to $h[R(j)]$ defined in Eq. (5), and discussed their efficiency related to the convergence of the interpolated PES: See Ref. 22.
[48] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[49] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).

[50] Since these reaction probabilities (88% and 75%) are estimated based on a small number (=64) of trajectories, they should be viewed within the context of the statistical uncertainty. Their standard deviations were calculated as 0.041 and 0.054, respectively.